Accepted Manuscript

Region based Image Steganalysis using Artificial Bee Colony

H. Sajedi, F. Ghareh Mohammadi

PII: S1047-3203(16)30251-6
DOI: http://dx.doi.org/10.1016/j.jvcir.2016.12.003
Reference: YJVCI 1907

To appear in: *J. Vis. Commun. Image R.*

Received Date: 2 August 2016
Revised Date: 25 October 2016
Accepted Date: 11 December 2016

Please cite this article as: H. Sajedi, F. Ghareh Mohammadi, Region based Image Steganalysis using Artificial Bee Colony, *J. Vis. Commun. Image R.* (2016), doi: http://dx.doi.org/10.1016/j.jvcir.2016.12.003

Region based Image Steganalysis using Artificial Bee Colony

H. Sajedi [1] and F. Ghareh Mohammadi [2]

[1]Dept. of Mathematics, Statistics and Computer Science, College of Science,
University of Tehran, Tehran, Iran
Email: hhsajedi@ut.ac.ir, Tel: +982161112915

[2]Faculty of Electrical and Computer Engineering, Tarbiat Modares University
Tehran, Iran;
Email: f.garemohammadi@modares.ac.ir

ABSTRACT

Steganalysis is the art and skill of discriminating stego images from cover images. Image steganalysis algorithms can be divided into two broad categories, specific and universal. In this paper, a novel universal image steganalysis algorithm is proposed which is called RISAB, Region based Image steganalysis using Artificial Bee colony. The goal of the proposed method is to realize a sub-image from stego and cover images through ABC with respect to density according to the cover, stego and difference images. In our method, we look for the best sub-image, which contains the highest density with respect to the changed embedding pixels. Furthermore, after selecting the best sub-image, we extract the features, which have been selected by IFAB, Image steganalysis based on Feature selection using Artificial Bee colony. At the end, both selected features by IFAB and extracted features by RISAB are combined. As a result, a feature vector is generated which improves accuracy of steganalysis. Experimental results show that our proposed method outperforms other approaches.

Keywords: Image Steganalysis; Artificial Bee Colony (ABC); Swarm Intelligent; Feature selection; Information hiding.

1. Introduction

In recent years, information hiding has been an active research area. Its application lies in Military and Intelligence agencies and counterintelligence. Early research has been focused on steganography to establish a secret channel between two parties. Nowadays, steganography and steganalysis are an active research area due to plenty of digital media serving as cover signals, and the availability of the public communication network as the internet. By secretly embedding information into an innocent cover object, transmitter hopes that the receiver will reach the message without suspicion.

Steganography is the skill of passing information through apparently innocent objects in a manner that the very presence of the message is unknown. The term steganography in Greek literally means, “Covered Writing”. The innocent objects can be referred to as cover text, cover image, or cover audio as appropriate. After embedding the message, it is referred to as stego text, stego image, or stego audio. A stego-key is used to control the embedding procedure, to restrict detection and/or recovery of the embedded data. While cryptography is about keeping the content of messages, steganography is about embedding the message, so that intermediate people cannot see the embedded message. For a more accurate behavior of the concept of steganographic safety, the person who reads refers to [1].

Steganalysis is the process of distinguishing the cover object from the stage objects. In fact, Steganalysis is the art and science of detecting messages hidden using steganography. The main goal of steganalysis is to detect the existence of secret hidden information in an object [2].

In recent decades, research in image steganalysis has been increased. Image steganalysis can be classified into two broad groups, namely, specific and universal approach. The specific steganalysis algorithms are considered with respect to steganography algorithms and focus on first analyzing the embedding process and next phase is to find some features of the cover image that become changed after finishing the embedding process. The design of specific steganalysis algorithms needs complete information of the steganographic algorithm. In other words, specific steganalysis essentially focus on the certain kind of steganography, which are used to embed messages. The latter one does not try to detect the certain hiding method (steganography), but causes correct identifications of detecting objects that may consist of hidden information; however, in comparison with specific steganalysis methods, blind steganalysis is more practical [3].

Lately, swarm intelligence has been thrilled significant interest in related fields either. The swarm intelligence is considered as the effort to redesign algorithms or widespread problem-solving strategies taken by the collective trait of social insect colonies and other animals being communities by Bonabeau et al.[4]. They determined their method on social insects lonely, like bees. The term swarm is taken in a general approach to state any limited collection of mutual agents or persons. The classical example of a swarm is bees swarming around their hive. Therefore, this example can easily be extended to other schemes with an identical structure [5, 6].

This paper proposes a new image steganalysis approach using artificial bee colony, in which feature selection and feature extraction are used, called Region based Image Steganalysis using Artificial Bee colony (RISAB). Its main purpose is to discover a sub-image which cause to increase the accuracy of detecting stego images from cover ones. The proposed method includes two separate phases: train phase and test phase. In the train phase, we have seven specific blocks according to the Fig. 1. In this phase, we implement Artificial Bee Colony (ABC) by applying density as fitness function to detect the best sub-image with length (L) and width (W), and then we verify and validate it by using energy. Finally, we create a learning model based on energy and selected features. Beside this, we have a test phase contains six independent blocks that based on the learned length (L), width (W), and the energy, we try to find the sub-image and extract features. By means of the most important sub-image, we try to extract the exact selected feature which are selected by IFAB[7]. The selected subset of features is the first half of the novel dataset and the extracted features by IFAB are the rest, so the generated dataset are able to discriminate the stego images from cover images properly. In summary, the goal of the proposed method is to realize a sub-image from stego and cover images through ABC with respect to density according to the cover, stego and difference images. The result declares, it is possible to observe that RISAB offers good results.

The content of the rest of the paper is structured as follows: some related and relevant works are presented and discussed in Section 2. The details of the proposed region based steganalysis method are presented in Section 3. Section 4 presents the experimental results. Finally, we have conclusion in Section 5.

2. Related Work

A universal steganalysis approach [8] most of the time adopts a learning based approach containing a training as well as a testing phase. In this process, a feature extraction phase is needed which is used in both training and testing phase. This feature extraction phase is employed to map an input image from a high-dimensional image space to a low-dimensional feature space. The training phase results a well-trained classifier. Out of many operative classifiers, Support Vector Machine (SVM), k-Nearest Neighbor (KNN), Neural Network (NN), etc., everyone can be selected. Decision boundaries are produced by the classifier to separate the feature space into positive sections (the stego image) and negative sections (the cover image) with the help of the created feature vectors, which are extracted from the training images. In the testing phase, with the help of the well-trained classifier with a specific decision boundary, an image can be categorized based on its feature vector’s domination in the feature space. If the feature vector classifies a section where the classifier is labeled as positive, the testing image is classified as a positive class. Otherwise, it is classified as a negative class. Therefore, according to the universal steganalysis, the number of common universal steganalysis features is described as follows.

Image Quality Features (IQM): most of the steganographic approaches may cause distortion in the stego image. The goal of image quality measures (IQMs) is providing a quantitative metric according to the image features for investigative this type of distortion. The statistical evidence produced by steganographic approaches may be recognized with respect to a set of IQMs and can be employed for detection as well [9]. With respect to search for a number of specific quality measures, which are consistent to steganographic distortions, the analysis of variance (ANOVA) algorithm can also be exploited. The effectiveness ranking of the quantitative metrics is applied based on the traditional F-*measure (or* F-score) which is generated by the ANOVA tests. These kinds of metrics can be applied as feature sets to discriminate between stego and cover images.

Calibration Based Features: Fridrich in [10] used both the calibration and the feature-based classification to plan a blind detector certainly to JPEG images. By extracting the features directly in the JPEG domain, it explains that the detection can be made more sensitive to a broader kind of embedding algorithms because the calibration process enhances the features' sensitivity to the embedding alterations when suppressing image-to-image variations.

Fridrich [10] proposed an approach of image steganalysis is well-known as FBS in which a set of distinctive features are gotten from DCT and spatial domains. As the key component of the proposed approach, is used to guesstimate statistics of the original image, before embedding, estimation is simply done by decompressing the JPEG image and then cropping its spatial representation by four lines of pixels in both horizontal and vertical directions. Later, the image is JPEG recompressed with the original quantization table. The deviation between statistics obtained from the given JPEG image and its original estimated version are attained via a set of functions that operate on both spatial and DCT domains.

Tavoli, et al [11] presented a feature weighting method for Document Image Retrieval System (DIRS) based on keyword spotting. They weighted the features using Weighted Principal Component Analysis (PCA). The purpose of PCA is to reduce the dimensionality of the data space to the smaller intrinsic dimensionality of feature space.

An algorithm based on image segmentation was proposed by Wang, et al [12], to utilize the content features of JPEG images. The images were segmented into a number of sub-images according to the texture complexity. The steganalysis features of each type of sub-images with the same or close texture complexity were extracted separately to build a classifier.

Moment Based Features: The impact of image steganography approach can be regarded as presenting some noise in the cover image. Some statistics of the image may be altered with the introduction of noise. These changes revealed heavily in the wavelet domain. From this concept Lyu and Farid [13] used the theory that the PDF of the wavelet sub-band coefficients and that of the prediction error of the sub-band coefficients will change after the process of data embedding. For instance, in a 3-level wavelet decomposition [14], the first four PDF moments, i.e., mean, variance, skewness, and kurtosis, of the sub-band coefficients at each high-pass orientation of each level are taken into attention as one feature set. The same kinds of PDF moments of the difference between the logarithm of the sub-band coefficients and the logarithm of the coefficients' cross sub-band linear predictions at each high-pass orientation of each level calculated may be considered as another features set. As a result, these two kinds of features produce acceptable results at high embedding rate.

Correlation Based Features: Local correlation pattern of an image might get disturbed after the process of data embedding is done. The inter-pixel dependency of a spatial image, and the intra-block or inter-block discrete cosine transform (DCT) coefficient dependency of a JPEG image correlation may be referred as correlation. Sullivan et al. [15] displayed the inter-pixel dependency by Markov chain and described it though a gray-level co-occurrence matrix (GLCM). In this context, the working principle of the three most widely used universal steganalysis algorithms namely BSM (Binary Similarity Measures)[16], WBS (Wavelet Based Steganalysis) [17] and FBS (Feature-Based Steganalysis) [10] has been mentioned.

Binary Similarity Measures (BSMs) has been established by Avcibas et al. [16] where extracted features are achieved from the spatial domain representation of the image. The authors supposed that correlation between the contiguous bit planes declines after a message is embedded in the image. In this case, the method looks at seventh and eighth bit planes of an image and computes three sorts of features, which include calculated similarity differences, histogram and entropy related features, and a set of measures with respect to a neighborhood-weighting mask.

Lyu and Farid [17] proposed a different approach in order to feature extraction from images called WBS. In their opinion most of the specific steganalysis algorithms focus on first order statistics, i.e., histogram of DCT coefficients, but simple countermeasures could keep the first-order statistics intact, which in turn makes the steganalysis algorithm passive. They proposed a model for natural images by using higher order statistics and displays that images with messages embedded in them differ considerably from this model.

3. Proposed method

There is a strong correlation between inter adjacent pixels. A good steganalysis tries to find a harmony which exists between inter adjacent pixels. By employing a steganography algorithm, the correlations between inter adjacent pixels are changed, so the harmony just is converted into other formats. In fact, there are some steganography algorithms, which embed data in image without regular embedding distribution. In other words, in some interested area of image, the rate of embedding is more than other areas.

According to the information discussed above, it is better to think about an idea of proposing an approach with respect to the transformed harmony. Therefore, we attempt to reduce the correlation between extracted features in order to increase the accuracy of steganalysis. In this study, we propose a Region based Image Steganalysis using Artificial Bee colony, RISAB. The whole structure of RISAB is given in Fig. 1 and 2 that include seven main separate phases and six main separate phases in training and testing process respectively.

The novelty of this study is to detect the best region, which met conditions. To that end, first of all, we find sample regions of images, then we evaluate them based on density and energy to choose the best one which is the best sub-image to extract features, which are selected from the original set by IFAB[7] , while the same features from whole the image are extracted. By union those two extracted feature sets, a new dataset for training phase are generated. It helps to facilitate the evolution of features for recognizing stego images from cover images by learning model. Furthermore, the novelty of RISAB is to extract distinctive features through which makes increase the accuracy of detecting images.

In this paper, Artificial Bee Colony (ABC) algorithm, firstly, is used in IFAB as a feature selection algorithm, then in RISAB, ABC is used for region selection by applying density [18] to evaluate generated sub-images. In ABC algorithm, the colony of artificial bees contains three groups of bees: employee, onlookers, and scout bees. A half of the colony stands for the employee artificial bees and the rest of it stands for the onlookers. One employee bee must exist per food source. It means, the quantity of employee bees is equivalent to the quantity of food sources. The employee bee of an abandoned food source turns into a scout bee.

Every cycle of the investigation contains three phases: changing the location of the employee and onlooker bees towards the food sources, determining their nectar quantities, and determining the scout bees. Therefore, the location of the scout bees is changed accidentally towards the available food sources. The food source shows possible solutions for a problem. The nectar amount of a food source corresponds to the value of the solution exposed by that food source.

Onlookers are placed on the food by employing ‘‘roulette wheel selection’’ method. The colony’s explorers surround scouts that have colony for each bee. The explorers have no guidance while looking for food. They are mostly worried about discovering any kind of food source; hence, the scouts are considered by low investigation costs, and a low mean in food source quality. Rarely, the scout can accidentally discover exclusively unidentified food sources. Regarding artificial bees, the artificial scouts can have the quick discovery of the class of potential solutions as a task. In ABC algorithm, the scout bee can be selected from one of the employee bees. The classification is controlled by a control element called ‘‘limit’’. If a solution representing a food source is not enhanced by a pre-arranged number of trials, the employee bee related to that food source turns into a scout. The number of trials for releasing a food source is equivalent to the value of ‘‘limit’’, that is a significant control factor of ABC algorithm [6].

The area density is also known as areal density, surface density, or superficial density, of a two-dimensional object is calculated as the mass per unit area. It can be calculated as Eq. (1):

$$D = \frac{m}{A} \qquad (1)$$

where D stands for average area density, m stands for total mass of the sub-image, and A stands for total area of the sub-image.

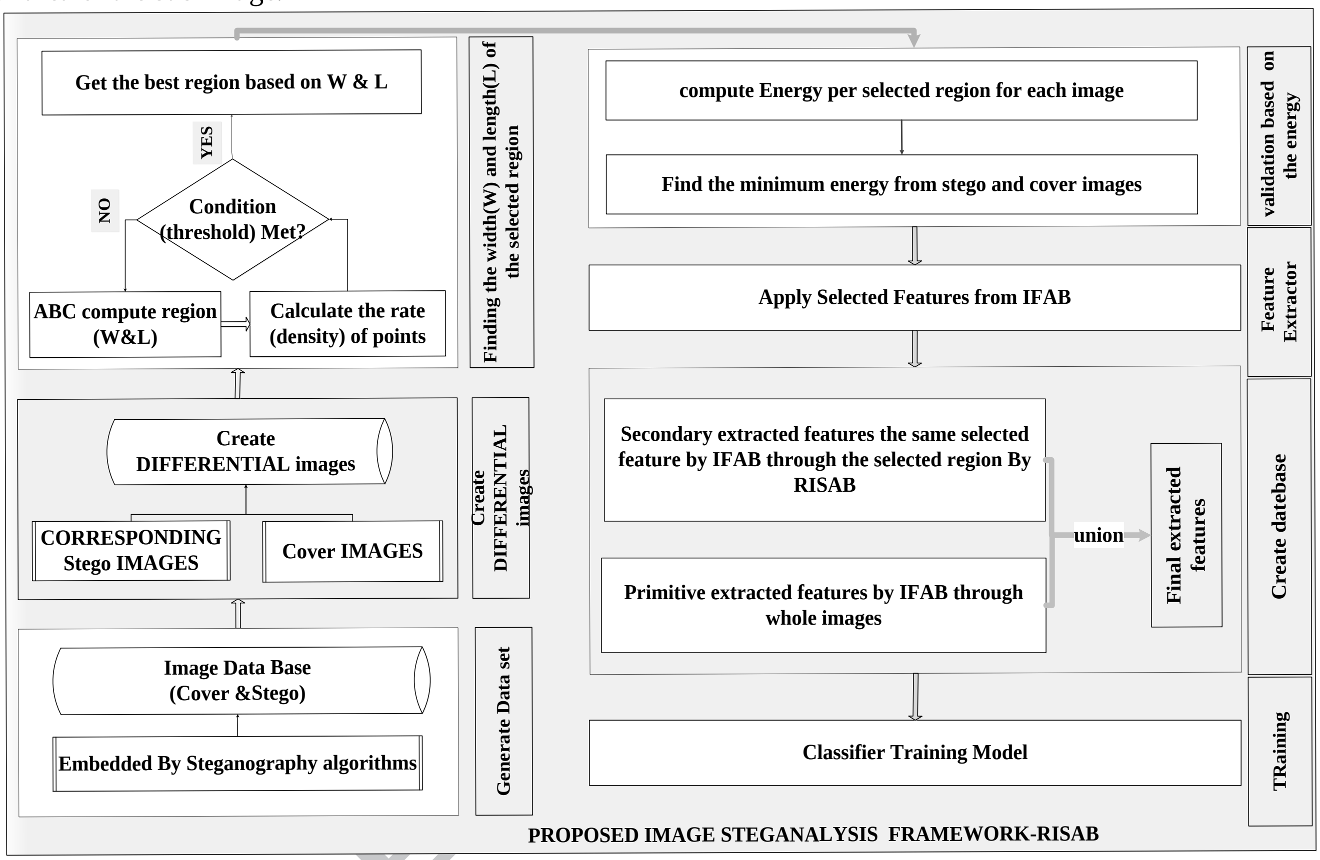

Figure. 1. The Block Diagram of the proposed method-RISAB (Training Process)

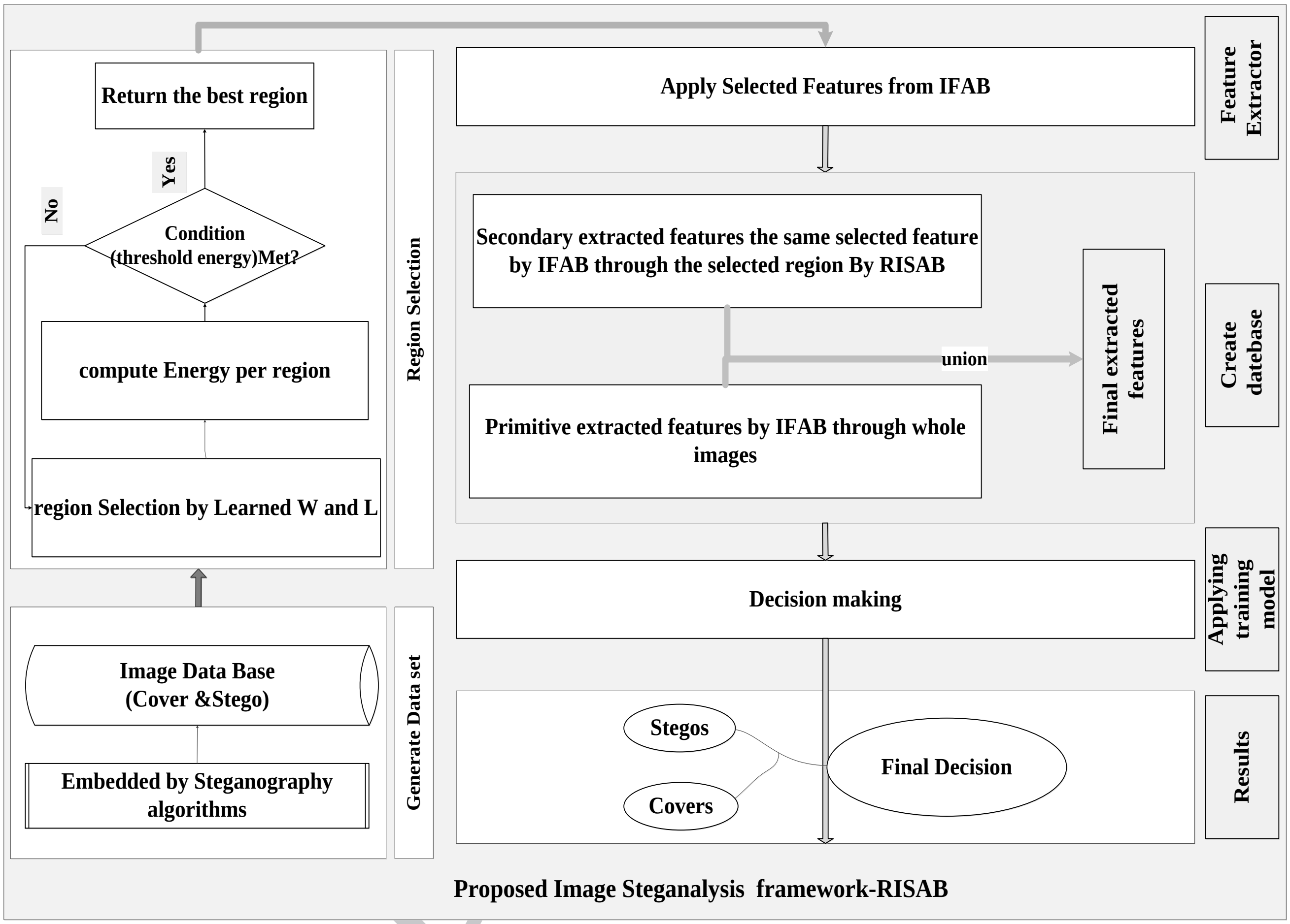

Figure 2. The Block Diagram of proposed method-RISAB (Testing Process)

3.1 Training Phase

According to Fig. 1, in which includes 7 blocks as different parts of training phase are described properly in what follows.

3.1.1 Generate Dataset of images (Applying a Steganography Method)

Gathering of reliable cover images and Generating data bases of stego images from corresponding images are constructed the first phase of the proposed method, and in this part, we employ a steganography algorithm in order to have complete both stego and cover image databases. We just use the cover images as input and employ HUGO[19] as a steganography algorithm to embed messages in images with an embedding ratio of 40% bpp. This phase of training process is the same with the testing process.

3.1.2 Create differential images (Identifying the embedding points)

By embedding a message in an image, the harmony of adjacent pixels in the image in some of its parts is altered. Finding the altered points in training process is one of the challenging issues in this approach. In this phase, we try to identify the length (L) and width (W) of a sub-image in which the points were changed by the embedding process. In order to handle this issue, we propose an approach which is given deeply bellow.

The approach uses both cover and stego images in order to identify the embedded points, which are obtained by using a well-known function given in Eq. (2). According to Eq. (2), two inputs are required. The first is stego image, which is the image within secret messages, and the second is the cover one, which is the image without messages. The result of Eq. (2) is an image includes the embedded points shown as white points and the rest are black. This subsection is skipped from the testing process because in testing phase we do not know which image is cover or stego.

$$EP = \sum_{i=1}^{M} \sum_{j=1}^{N} \; |s(i,j) - c(i,j)| \qquad (2)$$

where *EP* stands for Embedded Point, M and N are the length and width of the entrance image, which are standard with the same size, S stands for stego images, and C stands for cover images.

3.1.3 Finding the Width (W) and Length (L) of the best Region by employing ABC

This section presents a novel approach to resolve sub-image selection issues. The paper proposes two subsections to give a deeper explanation. Section 3-3-1 explains the schematic approach in details to select the best sub-image. The structure of the proposed ABC based sub-image selection is well explained in Section 3-3-2.

3.1.3.1 STRUCTURE OF THE PROPOSED SUB-IMAGE SELECTION APPROACH

Overall, a specific sub-image selection approach consists of four factors: a sub-image generation, an assessment function, a stopping condition in ABC and a sub-image validation process. The overall process of sub-image selection is evaluated by ABC, which is shown in Fig. 3. An important subject for the sub-image selection process is how to search the huge space of images. In training phase, based on each differential sub-image the algorithm finds a length and a width. The final L and W applicable in test phase, are obtained by averaging the Ls and Ws found in the training phase.

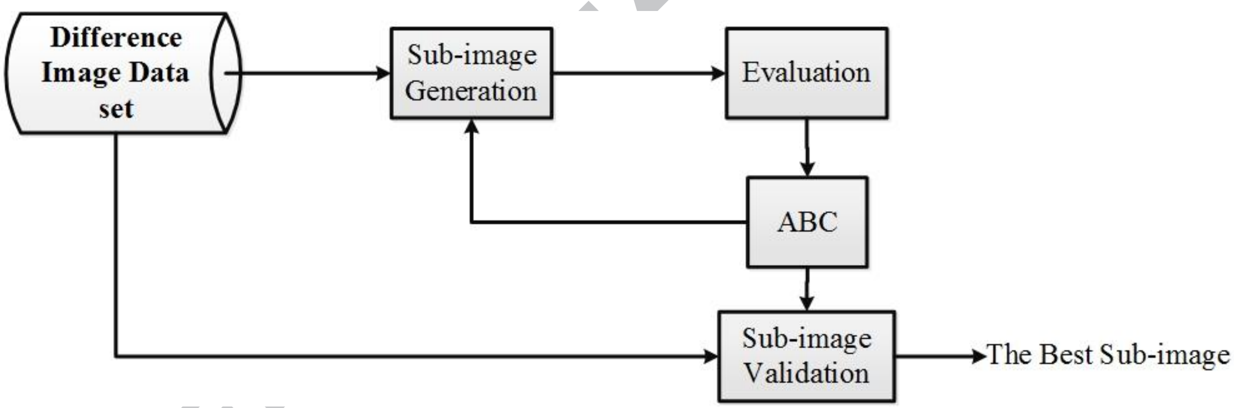

Figure 3. The general schematic approach to ABC-based sub-image selection

3.1.3.2 AN IMPROVED ARTIFICIAL BEE COLONY ALGORITHM FOR REGION SELECTION

In the proposed ABC based sub-image selection approach, ABC algorithm improves and optimizes the procedure of sub-image selection. Fig. 4 displays the pseudo code of implementing sub-image selection using ABC. Some parameters such as *M, N, Pi, Xi, Xj* are used where M stands for the height of the image, *N* stands for the width of the image, *Pi* stands for the chance to choose a solution, *Xi* stands for the Density of the sub-image is allocated to the employee bee and *Xj* stands for the Density of the sub-image the onlooker has chosen. ABC is employed as a sub-image selection approach and finds the best sub-image. A function is employed to evaluate every sub-image created by the onlookers, therefore, the suggested method is a kind of wrapper based methods. The steps of the proposed ABC based sub-image selection algorithm are given in Fig. 4 that is well explained as follows:

```
Begin
Cycle =1
Initialize ABC parameters
For each pixel in image (M × N)
```

```
    Evaluate the fitness of each individual Pixel using  Density
  Repeat
               // Employee Bees Phase
  Assign each generated sub-image to each employee bee.
  Produce new sub-images  Vi
  Pass the produced sub-image to the Density-based function
  Evaluate the fitness of the current solution using  Density
  Calculate the probability Pi of the current solution
               // Onlookers Bees Phase
   Select a new food source based on the probability pi
   Compute Vi  uses Xi and Xj
   Apply greedy selection between vi and xi
   Produce a food source
   Evaluate the fitness of the new solution using   Density
   If ( New Fitness > Current Fitness)
   Replace the new solution
               // Scout Bees Phase
   If (fitness of solutions does not improved )
   Abandon the current food source and try to find a new food source
      Memorize the best optimal sub-image
      Cycle = Cycle + 1
  Until  condition is met
  Employ the same searching procedure of bees to generate the optimal sub-image up to generation count
  End
```

Figure 4. The pseudo code of Density-based ABC method

3.1.3.3 INITIAL POPULATION

In this paper, the Density-based ABC approach is used to explore the novel search pixel space. Initial swarm is sometimes produced arbitrarily. With respect to generate an initial population with specific superiority and variety, a random approach is used to give us a perfect initial food source. Furthermore, the parameters of ABC including the number of food sources, the colony size, lower bound, upper bound, limit, max cycle and the maximum number of iterations are set. In fact, the population of employee bees and onlooker bees are equal to the dimension of covers or stego images. In addition, we use Density-based ABC algorithm. It evaluates the selected sub-images based on the density of embedded points in difference images.

3.1.3.4 COMPUTE AND EVALUATE PIXELS AND FOOD SOURCES

This section is related to evaluate the pixel and food source with respect to find the worth sub-images. It is important to evaluate both the pixel and food source because our proposed approach is a kind of wrapper based methods that helps to obtain the best sub-image. Density plays the main role of evaluation function in order to evaluate the candidate sub-images. The amount of density of the pixel and food source is used as an input into fitness function to evaluate the sub-images.

A feature vector has four dimensions, which includes the location of start point and end point of sub-images. The structure of food source is given in Fig. 5

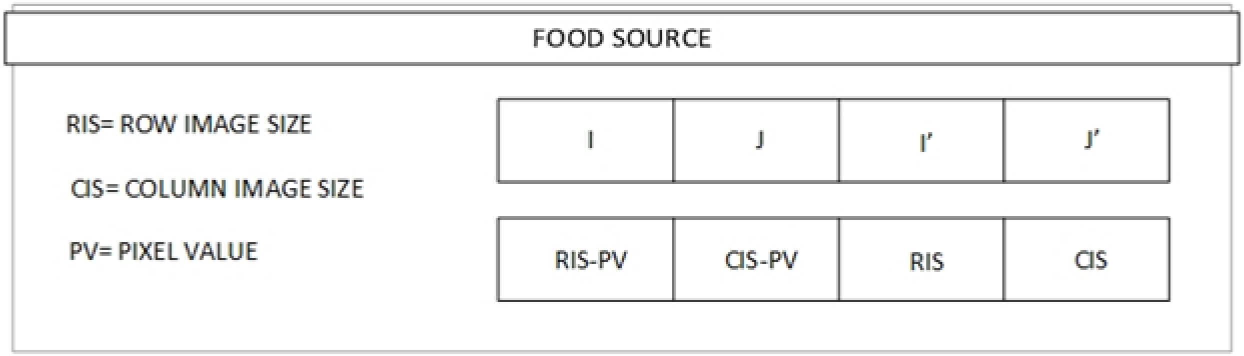

Figure 5. The food source format used in RISAB

According to the Fig. 5, I and J stand for the beginning point of sub-image and I', J' stand for the end point of the sub-image.

Each employee bee is assigned to each sub-image and evaluates the fitness of each feature by using Eq. (3). The novel candidate solution is then evaluated by means of a density-based function.

3.1.3.5 EMPLOYEE BEES

In this process of artificial bee colony algorithm, each employee bee makes food sources of their current positions and gets accuracy of its nectar fitness. In other words, each pixel in the food source is optimized based on the process of keeping informed feasible solution by employee bees. Employee bees are responsible for exploiting the food sources and share the information with the onlookers according to the nectar quality of the food sources which are exploited. The fitness of each solution is calculated by using Eq. (3). Where *D* stands for the amount of the sub-image density which comes from Eq. (1).

$$\text{Fitness (R)} = \begin{cases} \frac{1}{1+D} & D > 0 \\ 1+|D| & D < 0 \end{cases} \tag{3}$$

3.1.3.6 ONLOOKER BEES PROCESS

An onlooker bee picks up a highest solution fitness of food source regarding to its winning probability value. After sharing information with onlooker bees, performed by employee bees. The onlooker bees become aware of choosing a food source according to the probability of each food source using Eq. (4) and does the process of updating the current feasible solution based on Eq. (4) which is given below. Where m stands for the height of the images, s stands for the fitness of per solution (sub-image), and R stands for the fitness of selected solution by onlooker.

$$\mathrm{P(R)} = \frac{\mathrm{fitness(R)}}{\sum_{s=1}^{m} \mathrm{fitness(s)}} \tag{4}$$

After that, the onlooker bee analyzes the novel solution Vi based on the amount of sub-image density which the employee bee is pointing to and the sub-image the onlooker bee has chosen. The novel feasible solution Vi is allocated by Eq. (5).

$$j = rand\ [1, N] \quad V_i = f_i + \theta\,(f_i - f_j) \tag{5}$$

where i= {0,1,2…, N) and j= {0,1,2,…, N} and *N*= upper bound of pixel and *fi* stands for the density of the sub-image is allocated to the employee bee, *fj* stands for the Density of the sub-image the onlooker bee has chosen as a new sub-image candidate, and ɵ stands for a real random number in the range [-1, 1]. According to the Eq. (5), *Vi* is supposed to be the density of a new solution that has been selected by the onlooker Bee. If *Vi* < *fi* we can conclude that the density of *fj* is more than *Vi* , as a result of it, we have to change those *i, j* features together in order to enhance the total Density of food source. If $Vi > fi$ then we do nothing.

In this approach, every time the employee bee is allocated to a novel sub-image, the onlooker bee exploits to generate novel sub-images. After many possible sub-images have been exploited, the amount of nectar gets accumulated points to select a better cohesive sub-image. If none of employee bees have developed, after that, the employee bee will turn to a scout bee. The scout bee is allocated to a novel pixel space consisting in the Eq. (6). The recent solution in the onlooker bee's memory will be changed by the novel candidate solution if the novel solution's fitness is better than the current one.

$$X_i = X_{min} + \theta'(X_{max} - X_{min}) \tag{6}$$

where X_{max} and X_{min} stand for the upper and lower bounds of the number of population and θ' is a real random number in the range [0,1]. The bees keep performing the identical process until the stop condition met or best sub-image is prepared. The stop condition is that if the number of runs obtains the pre-determined maximum number, or if the amount of Density increments satisfied, the process finishes.

3.1.3.7 SCOUT BEES PROCESS

In this method, a scout bee generates a food source randomly in the predefined search pixel space. If the fitness value of the recent food source has not been boosted by predefined number of iterations, named the "limit", the food source will be left. Then, the scout bees will randomly generate a novel food source position in all pixel dimensions, which is calculated by using Eq. (6). This process is accomplished to avoid choosing the sub-optimal solution.

3.1.3.8 TERMINATION PROCESS

The process of bees will be continual until the number of runs gets the pre-determined maximum number, or satisfied the criteria.

3.1.4 Validation Based on Energy

According to the *L* and *W* which have been obtained in earlier phase, we try to check and verify sub-images based on the energy, and keep the image with respect to the best energy, based on the amount of it between both cover and stego images for whole testing data sets; the best energy is the minimum amount of energy between stego and cover images' energy. Fig. 6 illustrate a sample of (a) cover image, (b) corresponding stego image resulted by HUGO steganography method with payload of 0.4 and (c) difference image and a selected sub-image shown with a red rectangle. More details are presented in Section 3.2.2.

Cover Image	Stego image	Difference Image (red square is the selected sub-image in the train phase)

(a)	(b)	(c)

Figure 6. A sample of (a) cover image, (b) corresponding stego image resulted by HUGO steganography method with payload of 0.4, (c) difference image and a selected sub-image shown with a red rectangle.

3.1.5 Applying the selected feature set

According to our last proposed method called IFAB, which are discussed in 3.1.5.1 in general, the selected features are able to increase the predictive accuracy. Hence, by using IFAB, the best feature set has been selected, so with respect to the selected feature, we extract those from the specifically selected sub-images.

3.1.5.1 THE GENERAL STRUCTURE OF THE IFAB:

The general structure of the IFAB is given in Fig. 6 through which contains three important steps. These steps are the stages of the IFAB according to feature extractor, ABC feature selection, a classifier is called support vector machine (SVM).In IFAB, after finishing feature extractor step, we applied an approach to select the relevant and the most important features, called ABC, and then we applied SVM to evaluate the selected features set, and checked the result to validate, then if condition is met, the process will finish.

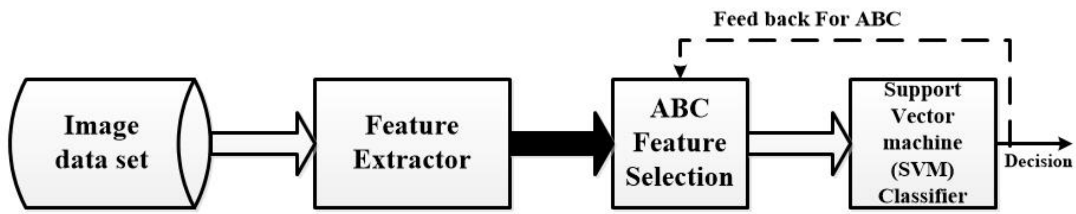

Figure 7. The structure of IFAB [7]

3.1.6 Create Database of extracted features (Generating a new dataset by steganalysis)

As we discussed above, after finishing the whole process, at the end we plan to generate a novel dataset based on those selected features by IFAB and those extracted features by RISAB. The generated dataset based on the steganalysis approach has different feature dimensions. In case of SPAM, the new dataset includes 160 features.

Finally, we add a class attribute, which is nominal, at the end of the attributes (features) to clarify that each instance relates to which kind of categories. This is done because of learning model of classification. If we do not import this attribute, the classifier cannot learn the model properly to predict in testing phase.

3.1.7 Learning model

This process is the last but not the least one. Upon creating the dataset from training images, it is time to learn a model, in which a classifier has to learn thoroughly the model of datasets. This phase is the important part of training phase, for it will be used in testing phase in order to detect the stego images from cover ones.

3.2 Testing Phase

By finishing the training phase, particularly the learning model, testing phase starts to evaluate images to recognize the stego images from cover ones. In the following, we describe different parts of testing phase, according to six important blocks shown in Fig. 2.

3.2.1 Generate Dataset (Applying a Steganography Method)

In this section, which is the first part of the proposed method in the testing process, we employ a steganography algorithm in order to have both complete stego and cover image databases. We just use the cover images as input and employ HUGO[19] as a steganography algorithm to embed messages in images with an embedding ratio of 40% bits per pixel (bpp). This phase of testing process is the same with the training process.

3.2.2 Region Selection

In the second block, which is the important part of testing phase, the goal is to determine the best region corresponding to the learned width (W) and length (L) of sub-images by implementing energy as Eq.(7). That is, for each input image, sub-images with W and L are selected and then the amount of energy per each is computed. Then, a sub-image is chosen as the best sub-image if its energy meets the threshold energy; otherwise, we try to select another sub-image until meet the threshold.

$$E = \sum_{r=0}^{L-1} P(r)^2 \qquad (7)$$

$$P(r) = \frac{n(r)}{L \times W}$$

where $L \times W$ represents the total number of pixels in the sub-image and $n(r)$ is the number of pixels of amplitude *r* in the same sub-image. L-1 is the maximum intensity of the sub-image.

After computing density and energy for each sub-image, we concluded that there is a direct relation between the energy of a sub-image in stego and cover images and density of a sub-image in the corresponding difference image. Therefore, in the training phase, while we are examining the density of points, we can calculate the energy of the correspond sub-image. Then, we selected a sub-image, which has high value not only in density but also in energy. Therefore, energy is a superior criterion in the case of determining which sub-image is well. Furthermore, in testing phase, as we are not able to obtain the density, thus, energy is used as the main criterion to select the best sub-image. Figure 8 shows a sample of (a) test image and (b) the selected sub-image shown with a red rectangle.

Figure 8. A sample of (a) test image, (b) the selected sub-image shown with a red rectangle.

3.2.3 Feature extractor

The third block of the proposed method relates to the feature extraction. This phase is to take the advantage of IFAB to lessen the amount of extracted features in regards to the increasing value which is calculated by detecting stego images from cover ones as accuracy. Based on the selected features by IFAB, which we have talked about it properly in the earlier section. We try to extract the exact the same features, from each of the selected sub-images. After that, the same features are extracted from whole images as well. This phase of testing process is the same with the training process.

3.2.4 Creating Dataset

In this section, the final database is created with combining two types of features. The first half of the features extracted from whole pixel of images and the second half of the features extracted from the corresponding selected sub-images. This phase of testing process is the same with the training process.

In testing phase, we do not import the class attribute to the created datasets. In fact, we try to anticipate and estimate how mature our method is. In other words, by testing phase we would like to determine the accuracy of RISAB.

3.2.5 Applying training model

After creating the dataset, we should use the learned model from training phase to anticipate the instances of testing dataset to check what percent of images are anticipated correctly or not. In this phase, we import each instance, which is in testing dataset, into the process of anticipating the label of instance using learned model of training phase.

3.2.6 Result

In the final section, all test images are labeled completely as stego and cover. Then, we realize that how many of them are true detecting and how many of them are false detecting.

4. EXPERIMENTAL RESULTS

To reveal the performance of the proposed region based method and to compare with other well-known approaches like MBEGA approach[20][21], two sets of experiments are carried out. For experimental results, the well-known steganography system (BOSS) version 1.01 was employed on grayscale image databases and the rate of embedding is 0.4 per pixel. This database contains 10000 cover images and 10000 stego images, so the number of classes in our experiments is binary. We have used a uniform random sampling by replacing approach to select 900 images out of 20000 images from BOSBASE site.

4.1 The feature extractors

In this paper, we applied both SPAM - Subtractive Pixel Adjacency Model [22] and CC-PEV[23][24] that are used to produce and extract the features for steganalysis. The length of SPAM 686 features consequently with the one extra feature called class in case of labeling the instances which including two types, the former is class ‘1’ stands for cover images and the latter is class ‘2’ stands for stego images.

4.2 Performance evaluation of proposed method

In this section, we examine the performance of the proposed method for different parameters settings. Since a sub-image with the size less than 160×160 pixels is too small for steganalysis, and extracting features is impossible, so we consider the size 160 as the minimum dimension of the sub-images by trial and error.

4.2.1 Parameter Tuning

4.2.1.1 - IFAB’s parameters

In this section, the performance of the proposed method for different parameters setting is investigated. The effect upon the subset features threshold is presented in Fig. 9 and Fig. 10. It is clear that the best amount for this parameter for SPAM is 80 and CC-PEV is 250. These numbers indicate that we should apply the subset features for the maximum possible predictive accuracy. According to these numbers, we can prove that our proposed method is capable of improving the performance of SPAM and CC-PEV algorithms.

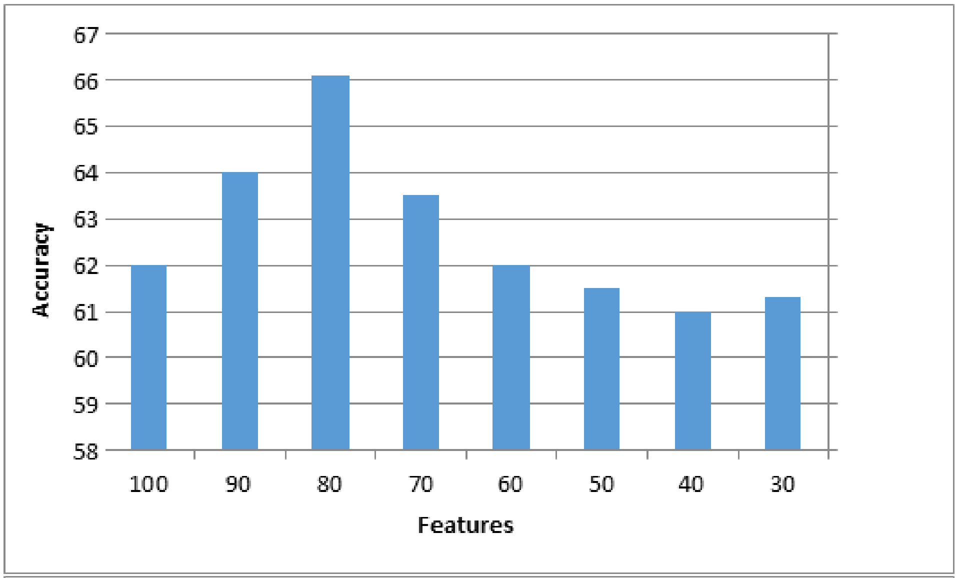

Figure 9. Different features subset effect on the accuracy of SPAM

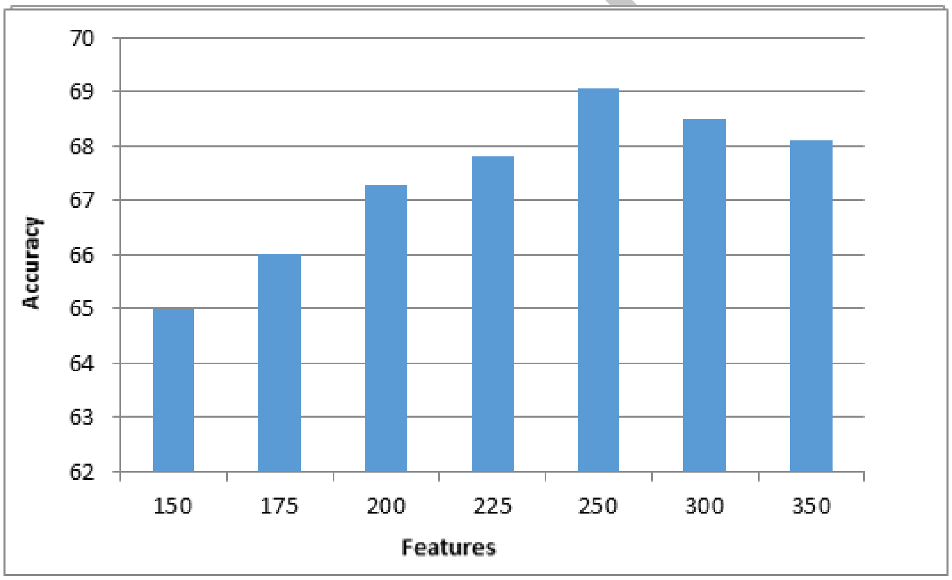

Figure 10. A comparison of the various kinds of features dimension on CC-PEV

Fig. 11 demonstrates how the food source values value effect on results in proposed method. It is obvious that by increasing the food sources, it increases the predictive accuracy significantly.

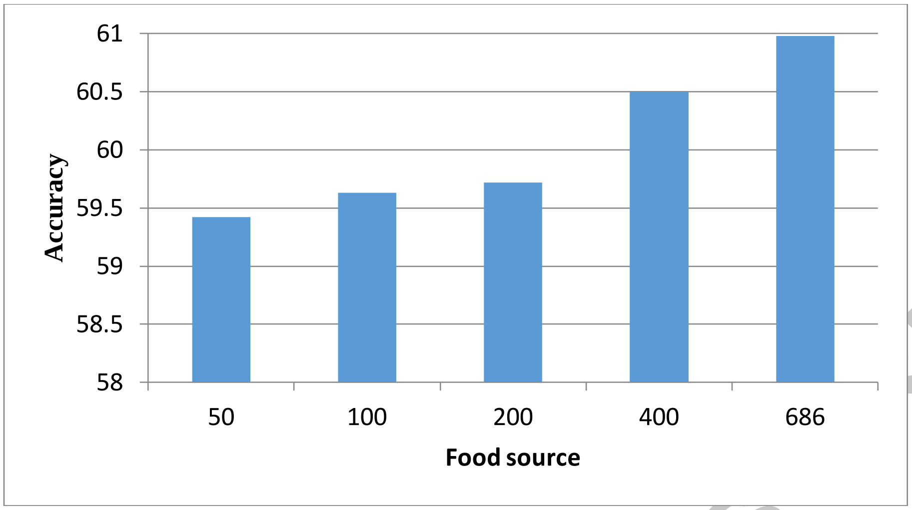

Figure 11. Food sources effect for the accuracy of spam

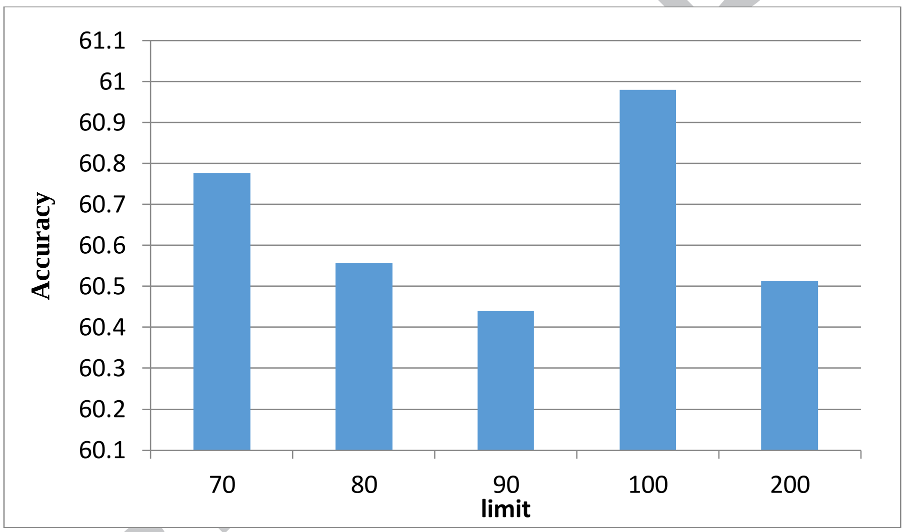

Figure 12. Different value of limits effect on the accuracy of spam

Fig. 12 reveals the effect on different kinds of limit value in proposed method. It is obvious that 100 is the ideal number to set the limit value. As mentioned, in every cycle may one of the employee bees turns to the scout bee, so it is needed control or handled by a control component called ‘‘limit’’. Then, the food source is abandoned by its employee bee, and the employee bee associated with that food source turns to a scout. The count of trials for releasing a food source is equivalent to the importance of ‘‘limit’’, that is an important control factor of ABC algorithm to prevent choosing the sub-optimal solution.

Artificial bee colony (ABC) has parameters less than other algorithms like ant colony optimization (ACO) which gives it a competitive edge over the other swarm intelligence algorithm. The initializing pre-defined parameters of IFAB corresponding to the Fig. 10, Fig. 11, and Fig. 12 are presented in Table 1

- RISAB’s parameters

In this subsection, we try to tune the ABC’s parameter to demonstrate the best region of candidate images. Food source is one of the important parameters. We conduct experiments with alternative values. Finally, we decide to assign a value which equal to the image size. In case of limit parameter, also the results show that the adequate limit value is 100. According to the parameters, the initializing pre-defined parameters of RISAB are presented in Table 1.

Table 1. Parameters specification in artificial bee colony

Parameter	Value(at RISAB)	Value(at IFAB)
Population size (P)	2*(image size)	2* Number of features in data set
Food source	P/2	P/2
Pixel Value (PV)	160	-
Feature Dimension (D)	4	80
Lower Bound	1	1
Upper Bound	N=548-PV	N
No. of runs	20	20
Limit	100	100

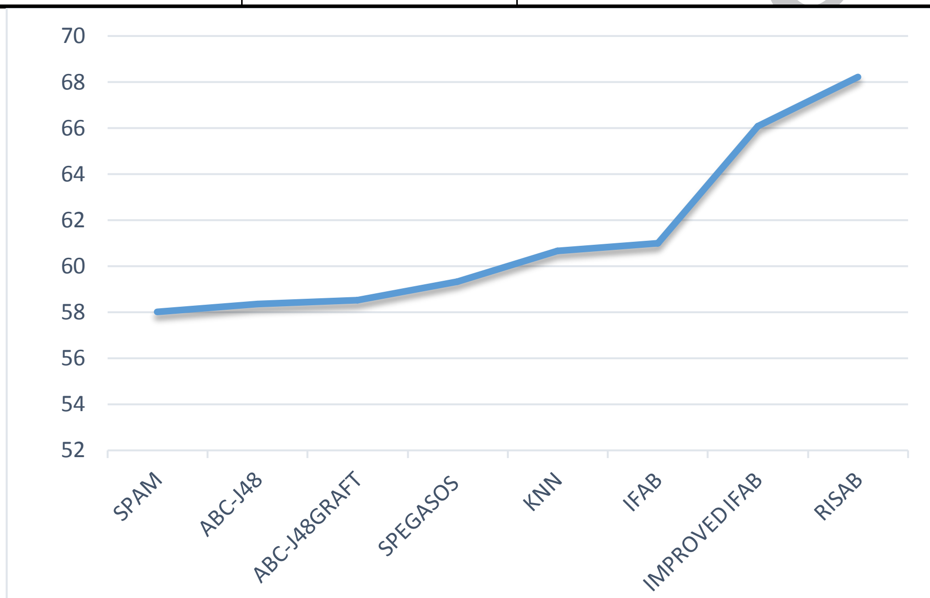

Figure 13. Comparing classification accuracy of different steganalysis methods based on features of SPAM

According to the Fig. 13, the SPAM data set is used, and it is obvious that the accuracy of our proposed method is higher than other methods. The accuracy of detecting cover images from stego images has improved almost 10 percent. Furthermore, Fig. 14 illustrates that RISAB is implemented on CC-Pev and the experimental results show that it does the best and outperforms other implemented methods. Although the accuracy of CC-Pev is better than SPAM, the rate of improved accuracy in case of SPAM is more than CC-Pev.

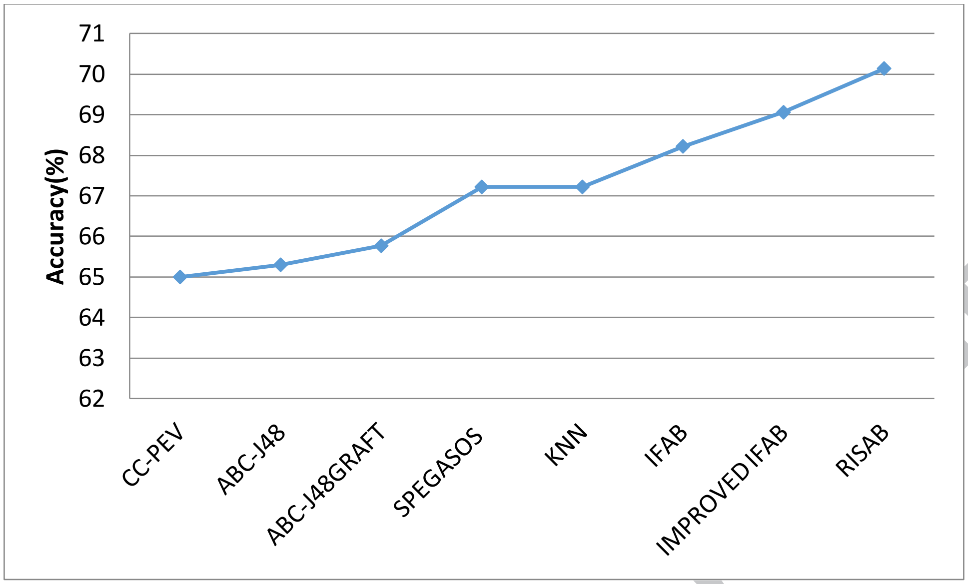

Figure 14. A comparison of the different steganalysis methods' classification accuracy on CC-Pev

4.3 Measurement

The measures used to evaluate the created model are as follows:

-*F-Measure* [25]

The weighted harmonic mean of precision and recall, in other words, it is the traditional F-measure or balanced F-score, which is obtained from Eq. (7).

$$F= \frac{2*precision*recall}{(precision + recall)} \quad (7)$$

-*Precision* [26]

Precision, which is mentioned in Eq. (8), is the fraction of the documents retrieved that are relevant to the user's information need.

$$precision = \frac{|\{relevant\ documents\} \cap \{retrived\ documents\}|}{|\{retrived\ documents\}|} \quad (8)$$

-*FP Rate*[27]

The proportion of non-relevant documents that are retrieved, out of all non-relevant documents available. Eq. (9) displays the proportion properly.

$$FPR = \frac{|\{non-relevant\ documents\} \cap \{retrived\ documents\}|}{|\{non-relevant\ documents\}|} \quad (9)$$

In binary classification, FPR is closely related to specificity (True Negative Rate) and is equal to (1-specificity). It can be looked at as the probability that a non-relevant document is retrieved by the query.

-*TP Rate (Recall or Sensitivity)* [28]

It measures the proportion of positives that are correctly identified.

-*Kappa coefficient* [29]

It is a statistical measure for qualitative (categorical) items regarding inter-rater agreement or inter-annotator agreement. Kappa statistic as a means for evaluating the prediction performance of classifiers. The kappa statistic measures the agreement of prediction with the completely positive class signifies complete agreement. The Eq. (10) is as follows:

$$k = \frac{p_o - p_e}{1 - p_e} \qquad (10)$$

where p_o stands for the relative observed agreement among raters, and p_e stands for the hypothetical probability of chance agreement, using the observed data to calculate the probabilities of each observer randomly saying each category. If the raters are in complete agreement then κ = 1.

-*Mean absolute error(MAE)*[30]

It is a risk function corresponding to the expected value of the squared error loss. In other words, The MAE measures the average scale of the errors in a set of predictions, without considering their direction. It measures accuracy for indiscrete variables. In statistics, the mean absolute error (MAE) is a quantity used to measure how close predictions are to the eventual outcomes. The Eq. (11) is given bellow.

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|f_i - y_i| = \frac{1}{n}\sum_{i=1}^{n}|e_i| \qquad (11)$$

As the name suggests, the mean absolute error is an average of the absolute errors $e_i = f_i - y_i$, where f_i stands for the prediction and y_i stands for the true value. The *MAE* is a linear score, which means that all the individual differences are weighted equally in the average.

-*Root mean squared error (RMSE)*[31]

It is a frequently used measure of the differences between values predicted by a model and the values actually observed. The RMSE is a quadratic scoring rule that measures the average scale of the error. Eq. (12) shows the way computing RMSE.

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n} {e_i}^2} \qquad (12)$$

-*Relative absolute error* [32]

It takes the total absolute error and normalizes it by dividing by the total absolute error of the simple predictor.

-*Root Relative Squared Error* [32]

By taking the square root of the relative squared error, one reduces the error to the same dimensions as the quantity being predicted.

Whole detailed accuracy of the proposed model by stego and cover class regarding true positive rate (TP), false positive rate (FP), precision, recall, f-measure, receiver operating character area (ROC Area) and precision recall curve area (PRC Area) is presented in Tables below.

-*Coverage of cases (0.95 level)*[33]

In order to determine the coverage of a case, it is necessary to define the retrieval and adaptation phases. A case solves a problem just if it is selected and retrieved by the similarity metric and if it is adapted to solve all features of the new problem.

4.4 RESULTS

RISAB is used to select the best sub-image to maximize classification performance, and minimize the number of extracted feature from 686 to 160 in case of SPAM, and from 548 to 500 in case of CC-PEV, with respect of increasing the accuracy of detecting stego images from cover ones. In order to compare RISAB with other methods to check how it outperforms other well-known methods, two separate and well-known datasets are chosen which are discussed as following.

According to the Table 2, firstly, the primary accuracy set of the SPAM is evaluated including some criteria such as F-measure, Recall, Precision, FP Rate, and TP Rate, and the results are shown is Table 2.

Table 2. Result of original SPAM

SPAM	TP Rate	FP Rate	Precision	Recall	F-Measure	Class
	0.604	0.508	0.551	0.604	0.576	cover
	0.492	0.396	0.546	0.492	0.518	stego
Weighted Avg.	0.549	0.493	0.549	0.549	0.547	

After getting SPAM accuracy, it is time to test IFAB to evaluate the selected features set, which has decreased the amount of feature to the 80 features. Table 3 can display proper view of results. It shows that having decreased the amount of features to exactly 80 features, IFAB increased the accuracy of SPAM. According to the Table 3, the amount of FP Rate decrease as other criteria increase.

Table 3. Result of selected Feature of SPAM by IFAB

IFAB	TP Rate	FP Rate	Precision	Recall	F-Measure	Class
	0.688	0.329	0.712	0.678	0.694	cover
	0.671	0.322	0.634	0.671	0.652	stego
Weighted Avg.	0.675	0.326	0.677	0.675	0.675	

After testing IFAB and SPAM, we evaluate RISAB to show how well is implemented. Table 4 reveals that the amount of FP Rate is the most minimum amount among SPAM primary accuracy, IFAB, and RISAB while others are the most amounts among them. As it shows, we get the top accuracy point for SPAM.

Table 4. Result of final features by RISAB

RISAB	TP Rate	FP Rate	Precision	Recall	F-Measure	Class
	0.713	0.282	0.752	0.713	0.732	cover
	0.718	0.287	0.676	0.718	0.696	stego
Weighted Avg.	0.715	0.284	0.717	0.715	0.716	

As a second dataset which include 548 features with primary accuracy of 47% is evaluated including some criteria such as F-measure, Recall, Precision, FP Rate, and TP Rate, therefore, based on the obtained test, the results are shown is Table 5.

Table 5. Result of original CC-PEV

CC_PEV	TP Rate	FP Rate	Precision	Recall	F-Measure	Class
	0.490	0.546	0.476	0.490	0.483	cover
	0.454	0.510	0.468	0.454	0.461	stego
Weighted Avg.	0.472	0.528	0.472	0.472	0.472	

Second, in order to check IFAB's performance based on the selected features, which has decreased the amount of feature to the 250 features. Table 6 can display proper view of results. It shows that by decreasing the amount of features to exactly 250 features, the accuracy increased. According to the Table 6 the amount of FP Rate decreases as other criteria increase.

Table 6. Result of selected Feature of CC-PEV by IFAB

IFAB	TP Rate	FP Rate	Precision	Recall	F-Measure	Class
	0.703	0.328	0.692	0.668	0.680	cover
	0.672	0.297	0.664	0.681	0.672	stego
Weighted Avg.	0.688	0.313	0.678	0.675	0.676	

Finally, by using RISAB, the best result has been reached. According to Table 7, the amount of precision goes to about 70 present. RISAB improves the amount of precision about 23 present. It also minimizes other criteria, those are revealed in Table 7.

Table 7. Result of final features of CC-PEV by RISAB

RISAB	TP Rate	FP Rate	Precision	Recall	F-Measure	Class
	0.713	0.302	0.702	0.703	0.702	cover
	0.698	0.287	0.686	0.698	0.692	stego
Weighted Avg.	0.706	0.295	0.694	0.701	0.697	

- Comparing SPAM and CC-PEV

Beside these criteria, there are some other criteria such as Kappa statistic, Mean absolute error (MAE), Root mean squared error (RMS), Relative absolute error, Coverage of cases (0.95 level). Table 8 reveals structure format, which SPAM, IFAB, RISAB all are in one table. Table 9 shows the results of CC_PEV, IFAB, and RISAB. High accuracy of RISAB shows that it outperforms the others.

Table 8. The result of comparing SPAM, IFAB, and RISAB.

Criteria	SPAM	IFAB	RISAB
Kappa statistic	9.62%	34.72%	42.85%
Mean absolute error (MAE)	45.11%	32.53%	28.48%
Root mean squared error (RMS)	67.16%	57.03%	53.37%
Relative absolute error	90.24%	65.59%	57.44%
Root relative squared error	134.34%	114.53%	107.18%
Coverage of cases (0.95 level)	54.88%	67.47%	71.51%

Table 9. The result of comparing CC_PEV, IFAB, and RISAB

Criteria	CC_PEV	IFAB	RISAB
Kappa statistic	5.5%	19.85%	24.49%
Mean absolute error (MAE)	52.78%	38.06%	33.32%
Root mean squared error (RMS)	72.65%	61.69%	57.73%
Relative absolute error	105.55%	76.71%	67.18%
Root relative squared error	145.29%	123.86%	115.91%
Coverage of cases (0.95 level)	47.22%	58.05%	61.52%

-Timing CC-PEV VS SPAM

Beyond the high accuracy, we need less time to detect stego images from the cover ones. By using RIAB, we could decrease the time of building the classifier model to discriminate cover and stego images. As Table 10 shows, we could decrease the build model time from 0.68s to 0.26s that is the essential criteria in this case. The required time for testing an image to investigate whether it is clean or stego is very small about milliseconds.

Table 10. Required Time for building models

	*Time primary	Time IFAB	Time RISAB
SPAM	0.68s	0.39s	0.26s
CC-Pev	0.96s	0.68s	0.62s
***Time taken to build model**			

Experimental results reveal that the proposed region based steganalysis (RISAB) method performs well for different images as compared to various other existing methods. It clearly indicates that the information-hiding rate has direct relation with the amount of density in that specific sub-image region. From the density analysis, it can be seen that the RISAB performs well in the presence of an amount of secret message in the selected region. This steganalyzer uses a SVM classifier to identify the cover and stego images, which produces the superiority of this method compared to the other existing ones

According to the definition of rich models like SRM [34,35], the overall goal is to capture a large number of different types of dependencies among neighboring pixels to give the model the ability to detect a wide spectrum of embedding algorithms. Thus, our proposed method is applicable. Selecting the best sub-image from each image could have positive impact of the accuracy of rich models.

5. Conclusion

In this paper, a novel region-based image steganalysis using an artificial bee colony algorithm is proposed, called RISAB. The goal of RISAB is to select a sub-image using artificial bee colony with respect to density according to the cover, stego, and difference image. By means of the most important sub-image, we would try to extract the exact selected features, which are selected by IFAB. The selected subset of features is the first half of the novel dataset and the extracted features by IRAB are the rest, so the generated dataset are able to discriminate the stego images from the cover images properly. Two methods are employed, namely SPAM and CC-PEV as steganalyzers. The performance of the proposed method has been compared with that some algorithms. Experimental results show that RISAB outperforms the mentioned algorithms. However, the main advantage of this paper is the strong specialized capability in recognizing the transformed harmony in part of images, and enhancing the performance in classification accuracy.

Acknowledgements

This research has been supported by a Grant from the INSF (No. 29999/01/01).

REFERENCES

[1] R. J. Anderson and Petitcolas, “On the Limits of Steganography.,” *IEEE Int.Journal. Communications*, vol. 16(4), pp. 474–481, 1998.

[2] F. G. Mohammadi and M. S. Abadeh, “A Survey of Data Mining Techniques for Steganalysis,” *Recent Advances In Steganography*, pp. 1–25, 2012.

[3] N. Fazio, A. R. Nicolosi, and I. M. Perera, “Broadcast steganography,” in *Topics in Cryptology–CT-RSA 2014*, Springer, pp. 64–84, 2014.

[4] E. Bonabeau, M. Dorigo, and G. Theraulaz, "Swarm intelligence: from natural to artificial systems." *Oxford University Press*, USA, 1999.

[5] D. Karaboga and B. Basturk, “On the performance of artificial bee colony (ABC) algorithm,” *Applied Soft Computing*, vol. 8, no. 1, pp. 687–697, 2008.

[6] F. G. Mohammadi and M. S. Abadeh, “A new metaheuristic feature subset selection approach for image steganalysis,” *Journal of Intelligent & Fuzzy Systems*, vol. 27, no. 3, 2014.

[7] F. G. Mohammadi and M. S. Abadeh, “Image steganalysis using a bee colony based feature selection algorithm,” *Engineering Applications of Artificial Intelligence*,vo.31, pp.35- 43, 2013.

[8] X.-Y. Luo, D.-S. Wang, P. Wang, and F.-L. Liu, “A review on blind detection for image steganography,” *Signal Processing*, vol. 88, no. 9, pp. 2138–2157, 2008.

[9] I. Avcibas, N. Memon, and B. Sankur, “Steganalysis using image quality metrics,” *Image Processing, IEEE Transactions on*, vol. 12, no. 2, pp. 221–229, 2003.

[10] J. Fridrich, “Feature-based steganalysis for JPEG images and its implications for future design of steganographic schemes,” in *Information Hiding*, pp. 67–81, 2005.

[11] R. Tavoli and E. Kozegar and M. Shojafar and H. Soleimani and Z. Pooranian "Weighted PCA for improving Document Image Retrieval System based on keyword spotting accuracy," *Telecommunications and Signal Processing (TSP),36th International Conference on. IEEE,*pp. 773 - 777, 2-4 July 2013.

[12] Wang, Ran and Xu, Mankun and Ping, Xijian and Zhang, Tao, " Steganalysis of JPEG images by block texture based segmentation,", Multimedia Tools and Applications,vol.74, no. 15, pp. 5725-5746 ,2015.

[13] S. Lyu and H. Farid, “Detecting hidden messages using higher-order statistics and support vector machines,” in *Information Hiding*, pp. 340–354, 2003.

[14] Y. Q. Shi, C. Chen, and W. Chen, “A Markov process based approach to effective attacking JPEG steganography,” in *Information Hiding*, pp. 249–264, 2007.

[15] K. Sullivan, U. Madhow, S. Chandrasekaran, and B. S. Manjunath, “Steganalysis for Markov cover data with applications to images,” *Information Forensics and Security, IEEE Transactions on*, vol. 1, no. 2, pp. 275–287, 2006.

[16] Avcibas, I.; Memon, N.; Sankur, B. “Image steganalysis with binary similarity measures,” *the mage Processing. Proceedings. International Conference on,* pp. 645– 648, 2002.

[17] H. Farid,“Detecting hidden messages using higher-order statistical models,” I*nternational Conference on Image Processing*, vol. 2, pp. II–905, 2002.

[18] N. S. Das and P. S. Rasmi, “Large-scale steganalysis using outlier detection method for image sharing application,” in *Circuit, Power and Computing Technologies (ICCPCT), International Conference on*, pp. 1–4, 2015.

[19] T. Filler and J. Fridrich, “Gibbs construction in steganography,” *Information Forensics and Security, IEEE Transactions on*, vol. 5, no. 4, pp. 705–720, 2010.

[20] Z. Zhu, Y.-S. Ong, and M. Dash, “Markov blanket-embedded genetic algorithm for gene selection,” *Pattern Recognition*, vol. 40, no. 11, pp. 3236–3248, 2007.

[21] S. Geetha and N. Kamaraj, “Optimized Image Steganalysis through Feature Selection using MBEGA,” *arXiv preprint arXiv:1008.2824*, 2010.

[22] T. Pevny, P. Bas, and J. Fridrich, “Steganalysis by subtractive pixel adjacency matrix,” *information Forensics and Security, IEEE Transactions on*, vol. 5, no. 2, pp. 215–224, 2010.

[23] F. Rodriguez, M. Di Martino, J. P. Kosut, F. Santomauro, F. Lecumberry, and A. Fernández, “Optimal and Linear F-Measure Classifiers Applied to Non-technical Losses Detection,” in *Progress in Pattern Recognition, Image Analysis, Computer Vision, and Applications*, Springer, 2015, pp. 83–91.

[24] J. Kodovsky, J. Fridrich, Calibration revisited, In J. Dittmann,S. Craver, and J. Fridrich, editors, Proceedings of the 11th ACM Multimedia and Security Workshop, Princeton, NJ, September 7–8, 2009.

[25] T. Pevny and J. Fridrich, Merging Markov and DCT features for multiclass JPEG steganalysis, In E. J. Delp and P. W. Wong, editors, Proceedings SPIE, Electronic Imaging, Security, Steganography, and Watermarking of Multimedia Contents IX, volume 6505, pages 3 1–3 14, San Jose, CA, January 29–February 1, 2007.

[26] J. Lu, F. Liu, and X. Luo, “A study on JPEG steganalytic features: Co-occurrence matrix vs. Markov transition probability matrix,” *Digital Investigation*, vol. 12, pp. 1–14, 2015.

[27] H. Zhang, X. Ping, M. Xu, and R. Wang, “Steganalysis by subtractive pixel adjacency matrix and dimensionality reduction,” *Science China Information Sciences*, vol. 57, no. 4, pp. 1–7, 2014.

[28] B. Akay and D. Karaboga, “A survey on the applications of artificial bee colony in signal, image, and video processing,” *Signal, Image and Video Processing*, vol. 9, no. 4, pp. 967–990, 2015.

[29] Z. Oplatkova, J. Holoska, M. Prochazka, R. Senkerik, and R. Jasek, “Optimization of artificial neural network structure in the case of steganalysis,” in *Handbook of Optimization*, Springer, pp. 821–843, 2013.

[30] X. Li, T. Zhang, Y. Zhang, W. Li, and X. Ping, “Quantitative steganalysis of spatial$\pm$1 steganography in JPEG decompressed images,” *Multimedia Tools and Applications*, vol. 73, no. 3, pp. 1487–1506, 2014.

[31] M. Goljan, J. Fridrich, R. Cogranne, and others, “Rich model for steganalysis of color images,” in *Parallel Computing Technologies (PARCOMPTECH), 2015 National Conference on*, pp. 185–190, 2015.

[32] S. Kamley, S. Jaloree, and R. S. Thakur, “Stock Market Behavior Prediction Using NN Based Model,” *British Journal of Mathematics & Computer Science*, vol. 4, no. 17, pp. 2502, 2014.

[33] D. Moldovan and S. Mutu, “Learning the Relationship Between Corporate Governance and Company Performance Using Data Mining,” in *Machine Learning and Data Mining in Pattern Recognition*, Springer, pp. 368–381, 2015.

[34] Kodovsky, J Fridrich, V Holub, Rich models for steganalysis of digital images, IEEE Transactions on Information Forensics and Security 7 (3), pp. 868-882, 2012.

[35]Wang P., Wei Z., Xiao L. Pure spatial rich model features for digital image steganalysis. Multimedia Tools and Applications, 75(5), pp. 2897-291, 2016.

Highlights

- Introducing a universal image steganalysis which is called RISAB
- Using Artificial Bee colony for region based Image steganalysis
- Region features and feature selection provide high detection accuracy